文章编号：1672-3961(2026)02-0019-16 DOI：10.6040/j.issn.1672-3961.0.2025.074

基于模块化网络的自适应加权联邦持续学习方法

周志刚[1]，孙博洋[1]，戴隆政[1]，白增亮[1]，苗钧重[2]
（1.山西财经大学信息学院，山西 太原 030006；2.哈尔滨工业大学网络空间安全学院，黑龙江 哈尔滨 150006）

摘要：针对资源受限环境下联邦持续学习（federated continual learning，FCL）中的横向与纵向灾难性遗忘问题，提出一种基于模块化网络的自适应加权联邦持续学习（modular-based adaptive weighted federated continual learning，MAWFCL）方法，有效应对任务演化引发的模型知识保持与任务适应性挑战。通过构建可组合的基础参数模块与自适应控制参数，实现个性化模型构建与任务适配；引入模块相似度度量机制，提升知识复用效率；结合参数容量感知的精准遗忘策略，有效控制模型复杂度；设计基于参数距离的自适应聚合算法，缓解聚合过程中的知识冲突。试验结果表明，MAWFCL 方法在准确率、灾难性遗忘抑制和通信效率方面优于现有方法，在 CIFAR100 数据集上的表现明显优于联邦生成重放学习（federated generative replay learning，FedGReL）方法和基于提示词的双重知识迁移（prompt-based dual knowledge transfer，Powder）方法，测试准确率分别提升10.93百分点和 10.17 百分点，在复杂任务中展现出显著优势。

关键词：联邦持续学习；参数隔离；神经网络；迁移学习；自适应算法
中图分类号：TP18 **文献标志码**：A
引用格式：周志刚，孙博洋，戴隆政，等. 基于模块化网络的自适应加权联邦持续学习方法[J]. 山东大学学报（工学版），2026，56(2)：19-34.
ZHOU Zhigang，SUN Boyang，DAI Longzheng，et al. Modular-based adaptive weighted federated continual learning method[J]. Journal of Shandong University（Engineering Science），2026，56(2)：19-34.

Modular-based adaptive weighted federated continual learning method

ZHOU Zhigang[1]，SUN Boyang[1]，DAI Longzheng[1]，BAI Zengliang[1]，MIAO Junzhong[2]
（1. School of Information，Shanxi University of Finance and Economics，Taiyuan 030006，Shanxi，China；2. School of Cyberspace Science，Harbin Institute of Technology，Harbin 150006，Heilongjiang，China）

Abstract：To address the challenges of horizontal and vertical catastrophic forgetting in resource-constrained federated continual learning（FCL）environments，a modular-based adaptive weighted federated continual learning（MAWFCL）method was proposed，which effectively addressed the difficulties of model knowledge retention and task adaptability caused by continuously evolving tasks. Personalized models were constructed by combining composable base parameter modules with adaptive control parameters to achieve adaptation to specific tasks. A module similarity-based reuse mechanism was introduced to enhance the efficiency of knowledge reuse. A parameter capacity-aware precision forgetting strategy was incorporated to prune low-contribution modules and maintain a compact model structure. An adaptive aggregation algorithm based on parameter distance was designed to alleviate knowledge conflicts during global model aggregation. Experimental results showed that MAWFCL method outperformed existing methods in terms of accuracy，catastrophic forgetting mitigation，and communication efficiency. On the CIFAR-100 dataset，MAWFCL method improved test accuracy over federated generative replay learning（FedGReL）and prompt-based dual knowledge transfer（Powder）by 10.93 percentage points and 10.17 percentage points，respectively，demonstrating significant advantages in complex tasks.

Keywords：federated continual learning；parameter isolation；neural network；transfer learning；adaptive algorithm

收稿日期：2025-05-23
基金项目：国家自然科学基金资助项目（61902226）；山西省自然科学研究资助项目（202403021221217，202203021221218）
第一作者简介：周志刚（1986— ），男，山西太原人，副教授，硕士生导师，博士，主要研究方向为联邦学习、云计算、隐私保护。
E-mail：zzgisgod@ sina.com

0 引言

持续学习是一种旨在应对异构任务流的学习范式[1]，智能系统在该范式下能够在学习未知任务的同时保持对历史任务性能的稳定[2]。在该范式启发下，智能系统正逐步从传统的静态学习模式向动态学习和适应性学习模式转变，形成通用人工智能（artificial general intelligence，AGI）[3]，展现出更接近人类学习行为的智能化水平。然而，仅依靠持续学习从直接经验中学习的方式是有限且低效的。为了进一步推动AGI的发展，个体系统还需要通过协同学习从其他系统中提取知识，实现跨任务迁移和高效适应复杂动态环境的关键目标[4]。联邦学习（federated learning，FL）作为一种分布式机器学习范式，允许终端设备在不传输本地数据的前提下，仅通过共享模型参数或梯度协同训练全局模型，有效解决客户端间为打破数据孤岛所面临的数据隐私泄露问题[5]。以联邦平均（federated averaging，FedAvg）算法[6]和联邦近端（federated proximal，FedProx）算法[7]为代表的联邦优化算法奠定了分布式协同训练的基础框架，为解决数据孤岛和隐私问题提供可行路径。

在联邦持续学习（federated continual learning，FCL）的多源协同框架下，智能系统的核心目标是在保障客户端数据隐私的前提下，通过分布式优化策略实现对多客户端私有异构任务流的联合学习[8]。由于FCL中的任务序列具有动态性，且客户端之间的数据分布往往存在非独立同分布性（Non-IID）[9]，两种特性在FCL中相互作用，使模型在局部训练与全局同步过程中更易出现知识偏移，进一步加剧模型的知识遗忘现象（称为灾难性遗忘问题），导致智能系统性能显著下降[10]。该问题已成为FCL领域研究的焦点，引起广泛关注和探究。文献[11]提出联邦加权客户端间传输（federated weighted inter-client transfer，FedWeIT）方法，将参数空间划分为多个独立的模块，为每个任务隔离专用的参数块，避免不同任务之间的干扰，有效减轻模型在FCL过程中的灾难性遗忘问题；文献[12]提出联邦个性化行为识别（federated human activity recognition，FedHAR）方法，针对设备资源受限与任务异构性共存的场景，采用硬件友好的二值掩码策略，将历史任务相关参数冻结，降低任务间冲突对模型性能的影响。这些研究极大提升了智能系统应对多领域动态任务流的学习能力，为智能系统在复杂场景中的应用提供强有力的支持。

尽管既有方法已取得一定成效，但它们在面对高度动态和异构的数据环境时仍能观察到性能衰退问题。一方面，这些方法可能依赖历史数据重现的强假设，在现实的动态演化中往往难以成立；另一方面，由客户端异构性引发的特征偏移会引入潜在干扰，进一步加剧模型训练难度和性能衰退。

面对客户端间的Non-IID数据分布，传统联邦聚合方法试图保留所有知识的策略往往适得其反，不仅未能有效整合异构信息，反而因无法剥离无关任务的干扰而损害全局模型的性能表现。近期研究（如联邦忘却学习相关工作[13]和精准遗忘联邦持续学习（accurate forgetting federated continual learning，AC-FCL）[14]）指出，遗忘现象并不总是有害的。AC-FCL证明精准知识遗忘策略可以通过有效去除无关或有害知识，减轻异构模型特征对全局模型的负面影响。目前，如何设计出一种精准知识遗忘机制以提升全局模型性能，依然是一个开放性问题。

客户端和服务器之间的通信是资源密集的，双方都面临大量的数据传输和存储压力。因此，在模型容量受限的前提下，如何降低通信成本并提高模型性能，仍然是FCL的关键挑战之一。

针对FCL在任务异构性、灾难性遗忘与资源受限环境下的挑战，本研究提出一种基于模块化网络的自适应加权联邦持续学习（modular-based adaptive weighted federated continual learning，MAWFCL）方法。该方法的核心思想是通过加性参数分解（additive parameter decomposition，APD）[15]和模块化网络结构[16]，将模型分解为基础知识参数和自适应控制参数，采用模块化网络对基础参数进行分块存储，实现多任务序列间的参数复用和自适应重整。

1 相关工作

1.1 持续学习

持续学习旨在使神经网络模型在吸收先前任务经验的基础上，高效适应后续任务并缓解灾难性遗忘问题。现有持续学习方法大致分为3类。第1类为基于正则化的方法，通过在损失函数中添加约束，限制新旧任务间重要参数的偏移。文献[17]提出弹性权重巩固（elastic weight consolidation，EWC）方法，基于Fisher信息矩阵约束参数变化；文献[18]提出无遗忘学习（learning without forgetting，

LwF)方法,采用知识蒸馏维持旧任务输出的一致性;文献[19]中的平均梯度情景记忆(averaged gradient episodic memory, A-GEM)方法利用梯度投影防止新任务学习破坏旧任务性能。第 2 类为参数隔离方法,通过结构性参数划分,动态扩展实现任务隔离与适应。文献[20]提出动态可扩展网络(dynamically expandable networks, DEN),通过迭代剪枝与分裂机制,动态扩展模型容量以适应任务变化;文献[21]中的专家门控网络引入任务感知门控机制,激活特定专家模块,结合结构扩展策略,减少任务干扰。第 3 类为基于重放的方法,通过保存或生成历史任务样本实现知识复用。文献[22]提出经验回放(experience replay, ER)方法,通过保存少量旧样本并与新数据混合重训练抑制灾难性遗忘;文献[23]提出暗经验回放(dark experience replay, DER),联合优化旧模型软概率与新任务损失,缓解新旧任务决策边界冲突;文献[24]提出增量分类器与表示学习(incremental classifier and representation learning, iCaRL)方法,结合样本回放与特征复习,避免对旧类别的遗忘。持续学习在监督学习、无监督学习与半监督学习任务中已得到广泛研究,涌现出多种具备实用潜力的策略。然而,将持续学习机制嵌入联邦学习框架,应对动态环境下 Non-IID 数据与多源任务异构性,仍属新兴课题,亟待深入研究与系统探索。

1.2　联邦持续学习

联邦持续学习主要围绕如何在多客户端异构环境下缓解灾难性遗忘与提升模型泛化能力展开研究。

为应对局部任务切换带来的知识遗忘,部分联邦持续学习研究聚焦局部持续训练策略,即鼓励各客户端在本地模型更新过程中保留历史知识。文献[25]提出联邦个性化混合表征(federated personalized mixture representation, FedPMR)方法,结合个性化表示与任务正则项提升旧任务适应性;文献[26]提出联邦类级自适应自蒸馏(federated class-wise adaptive self-distillation, FedCAD)方法,利用局部蒸馏从全局模型提取辅助知识,增强跨任务稳定性。这些方法能有效降低任务切换的遗忘风险,但仍局限于单客户端视角,普遍忽略了联邦聚合过程中因任务异构性带来的跨客户端干扰问题[27]。

一些研究试图通过增加适应性结构提升跨客户端的协同能力。文献[28]提出联邦类增量学习(federated class-incremental learning, FedCIL)方法,基于原型表示实现联邦类增量学习;文献[29]提出的梯度投影记忆(gradient projection memory, GPM)中引入梯度投影机制,限制关键梯度更新;文献[30]提出签名式任务知识整合的联邦持续学习(federated signature task knowledge integration, FedKNOW)方法,利用任务签名驱动知识整合,提升泛化能力;联邦曲率(federated curvature, FedCurv)方法[31]、联邦对比学习(federated contrastive learning, FedCL)方法[32]将基于重要性的参数约束扩展至联邦框架。上述策略多数仍聚焦在“保留”而非“选择性融合”视角,难以有效区分有助泛化的知识与易引发遗忘的干扰信息。

近年研究进一步引入提示词(prompt)学习与模块可控机制,提升联邦持续学习的泛化能力与表达灵活性。文献[33]的基于提示词的双重知识迁移(prompt-based dual knowledge transfer, Powder)方法通过 prompt 引导实现跨任务知识传递,显著提升任务适应性;文献[34]提出可追踪联邦持续学习(traceable federated continual learning, Traceable-FCL)方法,通过引入子模型标记与追溯结构,支持重复任务的检测与知识迁移;文献[35]结合原型增强与 prompt 激活策略,在 Non-IID 条件下显著提升系统稳定性;文献[36]提出多粒度提示个性化联邦持续学习(personalized federated continual learning via multi-granularity prompt, FedMGP)方法,通过融合多粒度提示机制引导模块化激活,缓解时空灾难性遗忘。这些方法在任务提示设计与子模块激活策略方面取得进展,但在面对任务差异显著、模型容量受限及客户端知识更新冲突等复杂场景时,仍难以实现灵活的结构控制和有价值知识的准确筛选。

鉴于上述问题,本研究所提 MAWFCL 方法尝试从以下 3 个维度系统性应对现有挑战:通过自适应聚合机制降低异构任务间的知识冲突,在全局模型聚合时有针对性地融合有助于泛化的知识;借助模块化结构进行参数划分与任务间知识隔离;设计精准遗忘机制,动态识别与移除对当前任务贡献较小的冗余模块,提升资源约束下的任务适应性。

2　问题定义

在标准的持续学习场景中,存在一系列流式任务序列 $T=\{T^1,T^2,\cdots,T^N\}$,其中 T^t 为第 t 个任务,$t=1,2,\cdots,N$。每个任务对应的数据集 $D^t=\{x_i^t,y_i^t\}_{i=1}^{M^t}$ 包含 M^t 对样本,其中 x_i^t 为第 t 个任务的第 i 个样本,y_i^t 为对应的标签。当学习第 t 个任务时,不能使用前 $t-1$ 个任务的数据。第 t 个任务的标签空间

为 Y^t，其中包含新的类别 C^t。这些类别与之前 $t-1$ 个任务中的既有类别不同，即 $C^t \cap (\cup_{j=1}^{t-1} C^j) = \varnothing$。本研究将传统的持续学习场景扩展到联邦持续学习场景中。给定 K 个本地客户端 $S=\{s_1, s_2, \cdots, s_K\}$ 和一个全局服务器 S_G，每个客户端 s_k 仅能访问其本地的任务序列 $T_k=\{T_k^1, T_k^2, \cdots, T_k^{N_k}\}$，其中 N_k 为客户端 s_k 的任务数量，T_k^t 表示客户端 s_k 的第 t 个任务。T_k^t 对应的数据集为 $D_k^t=\{x_k^{t,i}, y_k^{t,i}\}_{i=1}^{M_k^t}$，含有 M_k^t 对样本。在给定的训练任务 T_k^t 中，s_k 只能访问相应的数据集 D_k^t。

联邦持续训练共 R 轮全局聚合过程，在第 r（$r=1,2,\cdots,R$）轮聚合过程中将随机选择 n（$n<K$）个本地客户端 $S^r=\{s_1^r, s_2^r, \cdots, s_n^r\}$ 进行训练。训练开始前，服务器初始化全局模型参数 $\boldsymbol{\theta}_g^0$ 并分发至选中客户端。在第 r 轮聚合过程中，被选中客户端 s_k 将 $\boldsymbol{\theta}_g^{r-1}$ 初始化为本地基础知识参数 $\boldsymbol{W}_k$，并将 $\boldsymbol{W}_k$ 划分为 m 个模块，即 $\boldsymbol{W}_k=(\boldsymbol{W}_k^1 \quad \boldsymbol{W}_k^2 \quad \cdots \quad \boldsymbol{W}_k^m)$，其中 $\boldsymbol{W}_k^i$ 为客户端 s_k 的第 i 个基础参数模块。在训练任务 T_k^t 时，初始化自适应控制参数 $\boldsymbol{C}_k^t$，用于选择性激活参数模块；客户端的本地训练模型参数为 $\boldsymbol{\theta}_k^t$，$\boldsymbol{\theta}_k^t=\boldsymbol{W}_k \cdot \boldsymbol{C}_k^t$，在本地数据集 D_k^t 上进行训练，随后将损失函数 $L(\boldsymbol{\theta}_k^t, \boldsymbol{D}_k^t)$ 最小化，对 $\boldsymbol{W}_k$ 和 $\boldsymbol{C}_k^t$ 进行更新。当所有被选中的客户端完成任务训练后，将更新后的模型参数上传至服务器 S_G 进行全局聚合。S_G 根据客户端任务或模型间的相似性计算聚合权重并进行加权平均，生成新一轮全局模型参数 $\boldsymbol{\theta}_g^r$，供下一轮训练使用。联邦持续过程如图 1 所示。训练过程中需确保全局模型与本地客户端之间的高效协同，保证各客户端的本地训练与全局模型的聚合能够协调一致进行。

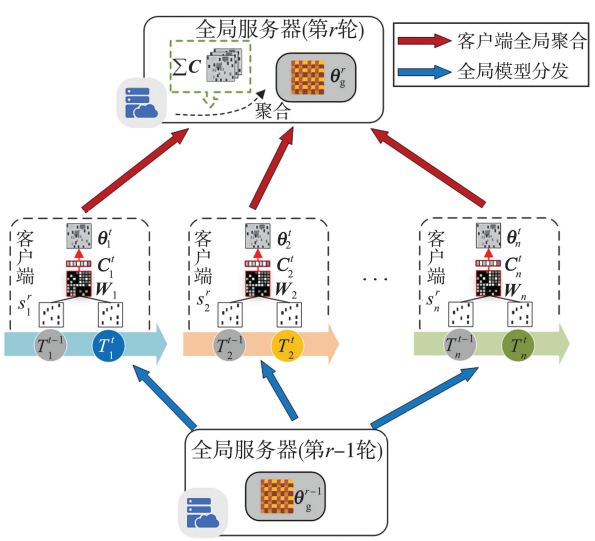

图 1　联邦持续过程

Fig.1　Federated continuation process

联邦持续学习旨在数据不可共享的前提下，通过多个客户端协同优化本地模型结构，使模型在学习当前任务的同时保留对历史任务的泛化能力。整体目标定义为

$$\min_{\{\boldsymbol{W}_k, \boldsymbol{C}_k^t\}} \sum_{k=1}^{K} \sum_{t=1}^{T} \left[L(\boldsymbol{W}_k \cdot \boldsymbol{C}_k^t, D_k^t) + \lambda R(\boldsymbol{W}_k, \boldsymbol{C}_k^t) \right], \tag{1}$$

式中：$L(\cdot)$ 为损失函数（如交叉熵损失）；λ 为正则化系数；$R(\cdot)$ 为正则化项，用于防止过拟合，增强模型泛化能力。

另外，在联邦持续学习场景中，模型需要在分布式环境下持续学习新任务，同时避免对历史任务知识的遗忘。为刻画模型在该过程中可能面临的不同遗忘机制，本研究提出横向灾难性遗忘和纵向灾难性遗忘的概念。

横向灾难性遗忘是指在单个客户端上，模型在顺序学习多个异构任务的过程中，由于学习新任务，历史任务知识被遗忘。令 s_k 持续接收任务序列 $T_k=\{T_k^1, T_k^2, \cdots, T_k^t\}$，若存在某一时刻 $j>i$，模型在完成任务 T_k^j 后的参数 $\boldsymbol{\theta}_k^j$ 在任务 T_k^i 上的性能低于训练后刚完成时的性能，即 $M(T_k^i; \boldsymbol{\theta}_k^j) < M(T_k^i; \boldsymbol{\theta}_k^i)$，则称该模型在任务 T_k^i 上发生横向灾难性遗忘，其中，$M(T_k^i; \cdot)$ 为 T_k^i 上的模型性能指标。横向灾难性遗忘现象会导致模型在历史任务上的性能下降，影响模型的稳定性与任务适应性。

纵向灾难性遗忘是指在全局模型聚合过程中，由于不同客户端之间的数据分布和任务存在显著的异构性，全局模型在聚合更新时丧失对部分客户端任务的适应性。全局模型在第 r 轮聚合后的参数为 $\boldsymbol{\theta}_g^r$，令 s_k 在完成本地任务 T_k^t 训练后模型参数为 $\boldsymbol{\theta}_k^t$，若满足 $M(T_k^t; \boldsymbol{\theta}_g^r) < M(T_k^t; \boldsymbol{\theta}_k^t)$，则称该模型在任务 T_k^t 上发生纵向灾难性遗忘。纵向灾难性遗忘现象会导致全局模型在某些客户端上的性能下降，影响模型的整体泛化能力和个性化适应能力。

3　模块化加权联邦持续学习方法

MAWFCL 模型采用模块化结构与加权策略相结合的方式进行构建，整体结构涵盖自适应知识重构、模块化稀疏复用、自适应平衡聚合算法及精准遗忘机制等关键部分，如图 2 所示。

MAWFCL 模型在运行过程中，自适应知识重构过程率先启动。在 FCL 的任务序列学习场景下，模型将获取的知识分别映射至本地基础参数和自适应控制参数。通过加性参数分解，将本地基础参

数精细划分为多个模块(如模块 1、模块 2 等)。这些模块各自存储不同侧重任务的通用知识,构成模型的知识储备基础。自适应控制参数依据当前任务需求,以二值化形式选择激活模块,动态重构出适应任务的模型参数。在模型训练阶段,客户端以最小化当前任务的局部损失函数为优化目标,结合对历史任务知识的保留机制,确保新旧任务间的有效迁移与融合,应对横向灾难性遗忘问题。

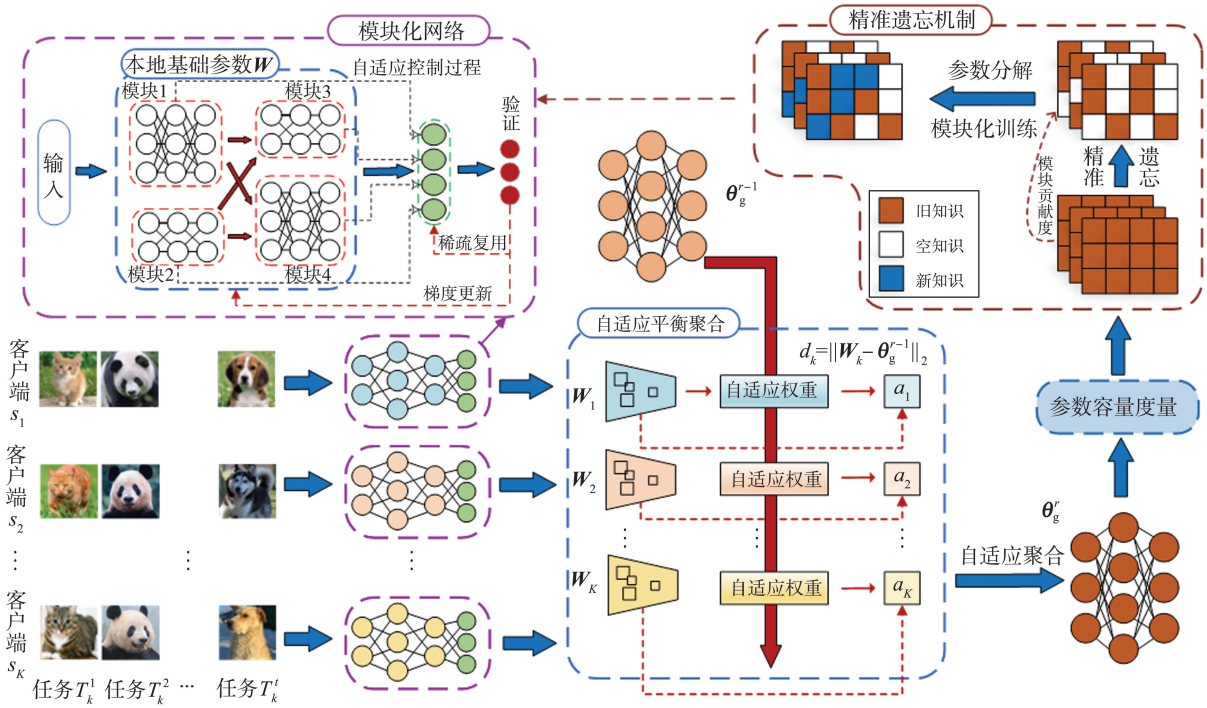

图 2　MAWFCL 模型结构图

Fig.2　Diagram of MAWFCL model architecture

在新任务学习进程中,MAWFCL 模型引入基于相似距离的复用度量机制,计算基础参数模块与新任务参数更新之间的相似距离。当相似距离小于预先设定的阈值时,模型将复用相应的基础参数模块。这一策略能够极大减少训练开销。此外,通过合理设计正则化项,模型能够保持各参数模块间的差异化,避免模块过度相似,在降低灾难性遗忘风险的同时,显著提高模型的复用能力。

当所有被选中客户端完成任务训练后,MAWFCL 模型进入自适应平衡聚合算法阶段。在服务器端,依据各客户端本地参数与全局参数之间的距离,动态分配自适应聚合权重。借助这些权重,服务器对客户端模型进行加权聚合,有效降低因客户端数据异构性引发的知识冲突,缓解纵向灾难性遗忘问题,保障全局模型在异构数据环境下的性能稳定。

随着任务数量不断增加,客户端本地的基础参数模块将不断扩展,可能导致参数容量达到上限。为了使模型能够继续学习新任务,确保性能不出现显著下降,MAWFCL 模型引入精准遗忘机制。该机制基于参数贡献度度量,从参数模块对历史任务的贡献度、在自适应控制参数中的使用次数及移除模块后对模型性能的影响等多个维度综合评估参数模块重要性,对已有的参数模块进行合理剪枝,实现对冗余或干扰特征的精准遗忘。

3.1　自适应知识重构过程

在联邦持续学习中,每一轮训练会涉及一组选中的客户端 $S^r=\{s_1^r,s_2^r,\cdots,s_n^r\}$,每个客户端 s_k 拥有自身的任务序列 $T_k=\{T_k^1,T_k^2,\cdots,T_k^{N_k}\}$,每个任务间存在显著的异构性和多样性。为缓解客户端在持续任务学习中可能出现的横向灾难性遗忘问题,本研究提出一种自适应知识重构的学习过程,支持基于任务需求的模型动态构建与参数选择。对于每个客户端 s_k,在训练过程中有 2 个核心参数:基础知识参数 $\boldsymbol{W}_k=(\boldsymbol{W}_k^1\quad \boldsymbol{W}_k^2\quad \cdots\quad \boldsymbol{W}_k^m)$ 和自适应控制参数 $\boldsymbol{C}_k^t=(c_k^{t,1}\quad c_k^{t,2}\quad \cdots\quad c_k^{t,m})$。其中,$c_k^{t,i}$ 为客户端 s_k 在任务 T_k^t 上的自适应控制参数中第 i 个模块参数,$c_k^{t,i}\in\{0,1\}$。在训练任务 T_k^t 时,客户端通过 $\boldsymbol{C}_k^t$ 对基础知识参数进行选择激活,构建任务模型参数

$$\boldsymbol{\theta}_k^t=\sum_{i=1}^{m}c_k^{t,i}\cdot\boldsymbol{W}_k^i,\tag{2}$$

式中,当 $c_k^{t,i}=1$ 时,表示激活第 i 个模块,当 $c_k^{t,i}=0$

时，表示不使用该模块。

为有效训练当前任务，保持参数稀疏性并控制模型复杂度，每个客户端的训练优化目标定义为

$$\min_{\{\boldsymbol{W}_k,\boldsymbol{C}_k^t\}} L(\boldsymbol{W}_k\cdot\boldsymbol{C}_k^t;D_k^t)+\lambda_1\|\boldsymbol{C}_k^t\|_1, \tag{3}$$

式中：$\|\boldsymbol{C}_k^t\|_1$ 为自适应控制参数的 L_1 范数，用于鼓励控制向量的稀疏性，即减少被激活的模块数量，降低通信成本；λ_1 为正则化系数，用以控制自适应控制参数的稀疏程度。在该过程中，$\boldsymbol{W}_k$ 通过模块化设计分解为多个可复用的子模块 $\boldsymbol{W}_k^i$，每个子模块对应该客户端的一部分历史通用知识。通过设计任务 $\boldsymbol{C}_k^t$ 实现对不同任务的个性化学习，客户端能够根据新任务的需求进行个性化调整，既可针对当前任务动态激活特定模块，以适配新任务需求，也可保持通用知识模块的稳定性，有效保留历史任务知识，缓解横向灾难性遗忘问题。

3.2 模块化稀疏复用策略

在自适应知识重构过程中，随着任务数量增加，为了进一步减少灾难性遗忘并提高参数复用能力，本研究设计模块化稀疏复用策略。

对 $\boldsymbol{W}_k$ 中各基础模块进行显式功能区域划分，将每个任务的知识映射到特定模块区域。每个任务新增的知识只更新与该任务相关的模块，不干扰其他模块，从而避免干扰其他任务参数。这种策略通过模块化分区，确保新任务与历史任务的知识在参数空间上的局部正交性，有效缓解跨任务学习中的灾难性遗忘问题。

客户端 s_k 学习新任务 T_k^t 时，通过训练得到新任务的参数更新 $\Delta\boldsymbol{W}_k^t$。为了提高模型的复用能力，确保复用的有效性，MAWFCL 引入一种基于相似距离的模块复用度量策略。在不同任务间，通过计算基础参数子模块 $\boldsymbol{W}_k^i$ 和参数更新 $\Delta\boldsymbol{W}_k^t$ 间的距离，判断它们之间的相似性，决定是否复用某些已有模块。定义相似距离

$$d_k^i=\|\Delta\boldsymbol{W}_k^t-\boldsymbol{W}_k^i\|_2, \tag{4}$$

式中 $\|\cdot\|_2$ 为 L_2 范数。通过 d_k^i 判别两个参数分布之间的相似性。如果存在某个模块 $\boldsymbol{W}_k^{i^*}$ 的相似距离 d_k^i 低于设定的阈值 τ，则表明 $\Delta\boldsymbol{W}_k^t$ 与 $\boldsymbol{W}_k^{i^*}$ 相似，可以复用该模块；若不满足此条件，则新建模块。基于余弦相似度函数的几何特性，当两个参数对应向量的夹角小于 41.4°时，它们的相似度将超过 0.75，表明它们在高维参数空间中具有较强的一致性。因此，本研究将 τ 设定为 0.75。该设定思路与现有模块共享结构学习方法研究中对模块相似度的判断标准基本一致[11]。该阈值的设定能够有效避免因无关模块误复用带来的负迁移问题，保留一定的结构共享空间，提升模型的参数利用率与泛化能力。

复用方式采用的参数累积更新策略为

$$\boldsymbol{W}_k^{i^*}\leftarrow\boldsymbol{W}_k^{i^*}+\eta\Delta\boldsymbol{W}_k^t, \tag{5}$$

式中，η 为学习率，用以控制更新程度。当客户端 s_k 学习新任务 T_k^t 时，为了防止模块间表示过于相似，降低模型表达能力，影响最终全局模型性能，MAWFCL 引入模块多样性正则化，模块间的差异性正则项

$$R_{\text{diversity}}=\sum_{i=1}^{m}\sum_{j\neq i}\|\boldsymbol{W}_k^i-\boldsymbol{W}_k^j\|_2^{-1}。 \tag{6}$$

$R_{\text{diversity}}$ 可以实现模块参数的多样性。根据上述策略，结合基础参数稀疏性控制，可以构建一个完整的模块化稀疏复用优化过程。客户端的最终优化目标为

$$\min_{\{\boldsymbol{W}_k,\boldsymbol{C}_k^t\}} L(\boldsymbol{W}_k\cdot\boldsymbol{C}_k^t;D_k^t)+\lambda_1\|\boldsymbol{C}_k^t\|_1+\lambda_2\sum_{i=1}^{m}d_k^i+\lambda_3R_{\text{diversity}}, \tag{7}$$

式中，λ_2、λ_3 为正则化系数。λ_2 控制模块复用度，鼓励模型优先复用与当前任务相似的已有模块，通过 d_k^i 度量任务间参数的可复用程度；λ_3 控制模块差异性正则项强度，鼓励模块保持足够的差异性。通过上述优化目标，MAWFCL 能够在最大化任务间的参数共享和复用的同时，保持模型的准确性和稀疏性。

3.3 自适应平衡聚合算法

在联邦学习场景下，各客户端的数据分布往往呈现出极为显著的异质性特征，即数据处于 Non-IID 状态。在进行全局模型聚合操作时，这一特性导致不同客户端之间极易引发知识冲突现象，使全局模型性能下滑。特别是在联邦持续学习环境中，每个客户端面临的任务序列各不相同，模型极易受到来自其他任务及客户端的干扰，由此产生纵向灾难性遗忘问题。鉴于此，本研究创新性地设计自适应平衡聚合算法，增强服务器端对客户端聚合贡献的动态感知能力，在知识融合过程中实现个性化控制与灾难性遗忘抑制。

在每一轮训练中，客户端 s_k 在当前任务 T_k^t 上完成本地训练，将更新后的 $\boldsymbol{W}_k$ 上传至服务器。服务器通过评估客户端的贡献实现动态加权策略，调整各个客户端在聚合过程中的权重，确保全局模型能够在融合各个客户端的知识时，最大限度地减少灾难性遗忘问题。服务器根据客户端参数的更新情况与历史任务表现计算自适应聚合权重

$$a_k = \frac{\exp(-\gamma d_k)}{\sum_{k=1}^{K} \exp(-\gamma d_k)}, \tag{8}$$

式中:γ 为调节参数,控制距离对权重分布的敏感性;$d_k = \| \boldsymbol{W}_k - \boldsymbol{\theta}_{\mathrm{g}}^{r-1} \|_2$,用于衡量客户端模型的相对重要性与贡献,$d_k$ 越小,表示模型更新方向越一致,权重越大。该设计使更新方向更一致、波动更小的客户端在聚合中获得更高权重,有效抑制异构更新的干扰影响,缓解纵向灾难性遗忘,提升模型在异构任务与长期训练过程中的稳健性与适应性。

服务器在收到每个客户端的本地基础知识参数 $\boldsymbol{W}_k$ 后,根据 a_k 对客户端的模型进行加权聚合,得到全局模型参数

$$\boldsymbol{\theta}_{\mathrm{g}}^{r} = \sum_{k=1}^{K} a_k \cdot \boldsymbol{W}_k 。 \tag{9}$$

3.4　模块网络优化目标

本研究结合自适应知识重构过程、模块化稀疏复用策略和自适应平衡聚合算法,构建一个统一的模块网络优化目标函数,应对联邦持续学习的任务异质性和灾难性遗忘问题等挑战。该优化函数融合局部任务损失最小化、模块稀疏性控制、模块差异性增强与聚合一致性正则化等多重目标,在保持模型性能的同时,实现结构压缩与知识保留的协同统一。

综合优化目标定义为

$$\min_{\{\boldsymbol{W}_k, \boldsymbol{C}_k^t\}} \sum_{k=1}^{K} \left\{ \sum_{t=1}^{|T_k|} \left[L(\boldsymbol{W}_k \cdot \boldsymbol{C}_k^t; D_k^t) + \lambda_1 \| \boldsymbol{C}_k^t \|_1 + \lambda_2 \sum_{i=1}^{m} d_k^i + \lambda_3 R_{\mathrm{diversity}} \right] + \lambda_4 d_k \right\}, \tag{10}$$

式中:$|T_k|$ 为客户端 s_k 已训练的历史任务数量;λ_4 为正则化系数,控制客户端与全局模型之间的一致性,通过 d_k 引导局部更新对齐全局目标。

通过上述优化设计,MAWFCL 能够在应对横向与纵向灾难性遗忘的同时,有效兼顾通信开销与计算效率,为联邦持续学习系统提供结构化、可解释、高性能的优化路径。

3.5　算法收敛性分析

为进一步验证 MAWFCL 算法在联邦持续学习框架下的优化可行性与训练稳定性,本研究从客户端本地训练与服务器全局聚合两个层面对算法的收敛性加以说明。

在本地训练层面,客户端针对当前任务构建的模型参数由特定基础知识模块组合而成,结构在任务内部固定。由于本地损失函数连续可导满足 Lipschitz 连续条件,采用梯度下降类优化方法进行迭代更新,局部训练过程可保证在适当步长下稳定收敛至局部最优解。模块选择与参数稀疏化过程受模块相似性阈值与容量限制的双重约束,使结构调整次数有限,有效控制任务切换时模型表示空间的扰动范围,维持学习过程的连续性与稳定性。

在全局聚合层面,MAWFCL 采用带权参数距离驱动的自适应聚合机制,本质为加权的联邦平均(FedAvg 拓展形式),收敛性已被广泛理论验证。该机制能够有效缓解因客户端间异构性导致的模型冲突,提升联邦优化的鲁棒性。结合客户端内部结构变更的可控性与服务器在每轮中对部分客户端的随机选择策略,整体训练过程在结构空间中的扰动被压缩在可收敛范围内。

综上所述,MAWFCL 在保障模块扩展性与参数共享能力的同时,能够在任务连续演化与资源受限环境下实现稳定迭代,训练过程具备良好的收敛性与泛化能力,为联邦持续学习提供可靠的优化保障。

4　精准遗忘机制与任务目标导向策略

4.1　参数容量度量

在联邦持续学习中,客户端会不断为新任务分配新的参数模块,提升模型对任务异构性的适应性。由于设备硬件资源有限,无法无限制地增加参数数量。因此,需要对本地基础参数容量进行度量和限制。设客户端 s_k 的最大参数容量为 $P_{\max}$,当前 $\boldsymbol{W}_k$ 的容量

$$P_k = \sum_{i=1}^{m} \Pi(\boldsymbol{W}_k^i), \tag{11}$$

式中 $\Pi(\boldsymbol{W}_k^i)$ 为参数模块 $\boldsymbol{W}_k^i$ 的参数量(如参数个数或占用的内存量)。当 $P_k \geqslant P_{\max}$ 时,触发遗忘机制。参数容量限制会影响模型的学习能力和性能,过大的容量会增加存储和计算成本,过小的容量可能导致模型无法充分学习。因此,本研究提出精准遗忘机制,在给定参数容量限制条件下,最大化模型的性能。

4.2　精准遗忘机制

当参数容量达到上限时,为了继续学习新任务,需要对已有的参数进行遗忘,同时保证模型性能不发生显著下降。精准遗忘机制旨在选择性地遗忘对当前和新任务影响较小的参数,避免模型整体性能显著下降。

为了实现精准遗忘,本研究提出参数模块贡献

度度量指标,结合参数模块对历史任务的贡献度、被调用频次及移除参数模块后对模型性能的影响进行综合评估。客户端 s_k 第 i 个参数模块的重要性

$$I_k^i=\alpha\cdot\frac{1}{T_k}\sum_{t=1}^{T_k}c_k^{t,i}+\beta\cdot\varphi_k^i, \tag{12}$$

式中:α 和 β 为权衡系数,用以满足不同的任务目标需求,$\alpha+\beta=1$; φ_k^i 为模块 $\boldsymbol{W}_k^i$ 参数置 0 后的模型性能下降程度,用于模拟参数模块移除效果。除参与模块重要性计算外,还要确保精准遗忘对模型性能的影响在可接受的范围内。如果性能下降超过阈值 $\delta_{\max}$(即 $\varphi_k^i\geqslant\delta_{\max}$),则停止遗忘或调整遗忘参数模块。基于模型在持续学习过程中性能波动的统计特性,当某一任务或模块在新任务训练后的准确率下降幅度超过 5%,通常可视为模型在该任务上的知识表达能力受到显著干扰[37]。因此,本研究将 $\delta_{\max}$ 设定为 0.05。$\delta_{\max}$ 能够有效区分正常性能波动与灾难性遗忘所致的干扰,避免过度响应,有效驱动对低贡献模块有控制地遗忘,实现参数空间释放与模型长期性能的平衡。

在每轮训练中,当 $P_k\geqslant P_{\max}$ 时,客户端会根据式(12)计算各参数模块的重要性,将参数模块按重要性从低到高排序,选择重要性最低的模块进行遗忘,更新 $\boldsymbol{W}_k$ 和 $\boldsymbol{C}_k^t$,直到满足 $P_k<P_{\max}$。

4.3 任务目标导向策略

本研究的主要任务目标为追求样本覆盖最大化和追求类别识别最大化。根据不同的任务需求,本研究提出两类遗忘度量,以增强模块筛选策略的定向性。

4.3.1 实例级遗忘度量

实例级遗忘度量旨在评估模型在学习新任务时对历史任务中具体实例的记忆情况。对于历史任务中的每一个样本实例,模型应尽可能保持识别能力,避免出现对实例的严重遗忘现象。实例级遗忘度量的目标是评估参数模块对具体实例的影响程度,以便在遗忘时优先保留对大量实例有贡献的参数,提高模型的泛化能力。参数模块 $\boldsymbol{W}_k^i$ 的实例级重要性度量

$$I_{\mathrm{dk}}^i=\alpha\cdot\frac{N_k^i}{N_k}+\beta\cdot\varphi_k^i, \tag{13}$$

式中 N_k^i 为参数模块 $\boldsymbol{W}_k^i$ 参与训练的样本数量。因为任务目前导向不同,全局模型聚合时,a_k 的计算也应根据任务需求进行相应调整。实例级遗忘度量下的聚合权重

$$a_{\mathrm{dk}}=\frac{N_k}{\sum_{i=1}^{K}N_i}, \tag{14}$$

式中 $\sum_{i=1}^{K}N_i$ 为全局客户端的总样本量。通过上述权重计算,引导系统优先保留对高覆盖样本群体有效的模块,避免对少数样本过拟合。

4.3.2 模型级遗忘度量

模型级遗忘度量关注模型在面对新任务类别时的表现,旨在评估模型在学习新任务时对新类别的适应性。随着新任务加入,模型不仅需要适应新类别的分类任务,还要避免被历史任务的知识干扰,拥有良好的泛化能力。模型级遗忘度量的目标是评估参数模块对整体模型性能的影响,特别关注对特定类别或任务的贡献,以便在遗忘时优先保留对模型性能关键的参数模块。参数模块 $\boldsymbol{W}_k^i$ 的模型级重要性度量

$$I_{\mathrm{ck}}^i=\alpha\cdot\frac{O_k^i}{O_k}+\beta\cdot\varphi_k^i, \tag{15}$$

式中,O_k^i 为参数模块 $\boldsymbol{W}_k^i$ 覆盖的类别数,O_k 为客户端 s_k 的总类别数。模型级遗忘度量下的聚合权重

$$a_{\mathrm{ck}}=\frac{O_k}{\sum_{i=1}^{K}O_i}, \tag{16}$$

式中 $\sum_{i=1}^{K}O_i$ 为全局客户端的总类别数。通过上述权重计算,鼓励模型保留对关键类别具有广泛贡献的结构模块,提升全局泛化性能。

4.4 综合优化目标函数

考虑参数容量限制和精准遗忘机制后,整体优化目标拓展为

$$\min_{\{\boldsymbol{W}_k,\boldsymbol{C}_k^t\}}\sum_{k=1}^{K}\left\{\sum_{t=1}^{|T_k|}\left[L(\boldsymbol{W}_k\cdot\boldsymbol{C}_k^t;D_k^t)+\lambda_1\|\boldsymbol{C}_k^t\|_1+\sum_{i=1}^{m}(\lambda_2 d_k^i+\lambda_5 I_k^i)+\lambda_3 R_{\mathrm{diversity}}\right]+\lambda_4 d_k\right\}, \tag{17}$$

式中,λ_5 为遗忘正则项权重系数,用于控制遗忘正则化对整体优化目标的影响程度。通过对被遗忘参数模块的重要性求和,在优化过程中惩罚遗忘重要模块的行为。在容量受限的情况下,确保模型在学习新任务时,不会因遗忘重要的历史知识导致性能大幅下降。

基于上述综合优化目标,MAWFCL 的整体训练流程如算法 1 所示。

算法 1 基于模块化网络的自适应加权联邦持续学习算法

输入 $S=\{s_1,s_2,\cdots,s_K\}$、$T_k=\{T_k^1,T_k^2,\cdots,T_k^{N_k}\}$、$\boldsymbol{\theta}_{\mathrm{g}}^0$、$R$、$P_{\max}$。

输出 联邦持续优化后的全局模型参数 $\boldsymbol{\theta}_{\mathrm{g}}^R$。

(1) 全局服务器初始化全局模型参数 $\boldsymbol{\theta}_{\mathrm{g}}^{0}$ 并将其划分为 m 个基础知识参数模块($\boldsymbol{\theta}_{\mathrm{g}}^{0,1}$ $\boldsymbol{\theta}_{\mathrm{g}}^{0,2}$ $\cdots$ $\boldsymbol{\theta}_{\mathrm{g}}^{0,m}$);

(2) 每个客户端 s_k 初始化本地基础知识参数 $\boldsymbol{W}_k=(\boldsymbol{W}_k^1\ \ \boldsymbol{W}_k^2\ \ \cdots\ \ \boldsymbol{W}_k^m)=\boldsymbol{\theta}_{\mathrm{g}}^{0}$;

(3) for 全局轮次 $r=1:R$;

(4) 服务器随机选择 n 个客户端 $S^r\subseteq S$;

(5) for 每个客户端$s_k\in S^r$;

(6) 根据当前任务 T_k^t 初始化自适应控制参数 $\boldsymbol{C}_k^t$,根据式(2)构建任务模型参数 $\boldsymbol{\theta}_k^t$,对 $\boldsymbol{\theta}_k^t$ 进行训练,根据式(3)最小化本地损失函数,实现稀疏保存;

(7) 根据式(4)计算新任务的参数更新与基础模块的相似距离 d_k^i,依据设定的阈值 τ 选择复用相似模块 $\boldsymbol{W}_k^{i^*}$;

(8) 根据式(5)更新客户端模型参数;

(9) 若客户端 s_k 的参数量 $P_k\geqslant P_{\max}$,根据式(12)计算 I_k^i,移除其中价值最低的模块;

(10) 将更新后的本地参数 $\boldsymbol{W}_k$ 上传至服务器;

(11) end for;

(12) 根据式(8)计算客户端权重 a_k,根据式(9)进行全局聚合,生成新一轮全局模型参数 $\boldsymbol{\theta}_{\mathrm{g}}^{r}$;

(13) end for;

(14) 返回 $\boldsymbol{\theta}_{\mathrm{g}}^{R}$。

5　试验及结果分析

为验证 MAWFCL 模型在联邦持续学习场景中的有效性,本研究设计并实施多组实证评估试验,涵盖模型性能、灾难性遗忘抑制能力、模块机制作用及通信效率 4 个维度,分别从试验设置、对比试验、消融试验与通信效率评估等方面进行分析。

5.1　试验设置

5.1.1　数据集

本试验使用 5 个经典图像分类数据集(MNIST、FashionMNIST、SVHN、CIFAR10 和 CIFAR100)进行评估,以确保本研究模型能够适应不同类型和难度的数据集,测试模型在异构环境中的性能。MNIST 数据集有 10 个类别,包含 60 000 张手写数字(涵盖 0~9)的训练图像和 10 000 张测试图像,所有图像为灰度图像,尺寸为 28 像素×28 像素,是图像分类领域中经典的基准数据集之一。FashionMNIST 数据集同样有 10 个类别,涵盖不同种类的服装物品(如 T 恤、裤子、鞋子等),包含 60 000张训练图像和 10 000 张测试图像,所有图像为灰度图像,尺寸为 28 像素×28 像素,用于替代 MNIST 数据集进行图像分类算法的评估,且分类难度较高。SVHN 数据集包含由 Google 街景图像中提取出的房屋号码,包括 73 257 张训练图像和 26 032张测试图像,所有图像为彩色图像,尺寸为 32 像素×32 像素,与 MNIST 数据集类似,但背景复杂,包含更多噪声,因而更具挑战性。CIFAR10 数据集包含 60 000 张 32 像素×32 像素的彩色图像,分为 10 个类别,每类 6 000 张图像,广泛用于图像分类和深度学习模型的评估,图像复杂度较高,类别丰富,在评估模型的泛化能力和鲁棒性方面具有较高的参考价值。CIFAR100 数据集包含 60 000 张 32 像素×32 像素的彩色图像,分为 100 个类别,每类 600 张图像,与 CIFAR10 数据集相比有更高的难度,包含类别更多,广泛应用于图像分类和深度学习的研究,适合测试模型在多类别分类任务中的性能和泛化能力。

5.1.2　对比模型

本试验选取 6 种联邦学习模型(FedAvg、FedCurv、FedHAR、FedProx、联邦生成重放学习(federated generative replay learning, FedGReL)、Powder)作为对比模型,每个模型采用不同的聚合策略和优化方法,旨在探究它们在不同数据集上的表现。FedAvg 为联邦学习基准方法,通过简单平均客户端参数更新全局模型,但缺乏持续学习机制,用于验证朴素方法的局限性。FedCurv 为基于 EWC 的联邦持续学习方法,通过计算参数 Fisher 信息矩阵,约束重要参数的变化幅度,保护历史知识,适用于简单任务,对复杂任务的计算开销较大。FedHAR 为硬件友好型参数掩码方法,通过冻结历史任务相关参数区域隔离知识冲突,通过二值掩码降低计算成本,但固定掩码策略可能限制模型容量。FedProx 为针对异构数据设计的联邦优化方法,在本地目标函数中添加近端项,约束客户端模型与全局模型的偏离,可缓解 Non-IID 数据分布问题,但未显式处理持续学习中的遗忘。FedGReL 为基于生成对抗重放的联邦持续学习方法,客户端训练生成器模拟历史数据分布,避免存储真实数据,生成器与分类器交替训练,在隐私保护场景下具有优势,但对计算资源需求较高。Powder 为基于 prompt 学习的联邦持续学习方法,采用双重知识传递机制,通过引导模型在多任务之间传递知识,显著提升任务适应性,能够在复杂异构环境下快速适应新任务,减少模型训练成本,提高跨任务知识的

有效迁移。本研究将FedAvg作为基准模型应用于所有数据集，FedCurv和FedHAR模型应用于MNIST和FashionMNIST数据集，FedProx和FedGReL模型应用于SVHN和CIFAR10数据集，Powder模型应用于所有数据集并与其他模型进行比较。

5.1.3 评判指标

（1）任务准确率A_{T}：模型在任务测试集上的Top-1分类准确率，反映模型对新任务的学习能力。

（2）损失L_{oss}：在训练过程中，每个训练轮次结束后记录的损失，是评估模型训练效果的重要依据。L_{oss}越低，表明模型的拟合程度越优。

（3）收敛速度S_{C}：衡量训练过程中模型达到稳定准确度所需的通信轮次。训练轮数越少，意味着模型S_{C}越快。

（4）横向灾难性遗忘率r_{HCF}：用于评估同一模型在不同训练阶段针对同一训练任务的性能衰退情况。r_{HCF}越小，表明灾难性遗忘程度越低，即模型的性能衰退越不明显。r_{HCF}的计算式为

$$r_{\mathrm{HCF}}=\frac{1}{|T_{\mathrm{c}}|}\sum_{t=1}^{|T_{\mathrm{c}}|}\left(\frac{A_{\mathrm{Tc}}(t\to t)-A_{\mathrm{Tc}}(|T_{\mathrm{c}}|\to t)}{A_{\mathrm{Tc}}(t\to t)}\right),$$

式中：$|T_{\mathrm{c}}|$为本地模型的已学习任务总数；$A_{\mathrm{Tc}}(t\to t)$为客户端在完成第t个训练任务后，其本地模型在第t个任务测试集上的任务准确率；$A_{\mathrm{Tc}}(|T_{\mathrm{c}}|\to t)$为在完成所有任务训练后，客户端的本地模型在第$t$个任务测试集上的任务准确率。

（5）纵向灾难性遗忘率r_{VCF}：用于衡量全局模型在各客户端本地任务中的性能差异程度。通过计算所有客户端任务准确率的标准差，体现模型在异构环境下的泛化一致性。r_{VCF}越小，表明全局模型在异构环境下的泛化一致性越好，灾难性遗忘程度越低。r_{VCF}的计算式为

$$r_{\mathrm{VCF}}=\frac{1}{|T_{\mathrm{g}}|}\sum_{t=1}^{|T_{\mathrm{g}}|}\frac{A_{\mathrm{Tg}}(t\to t)-A_{\mathrm{Tg}}(|T_{\mathrm{g}}|\to t)}{A_{\mathrm{Tg}}(t\to t)},$$

式中：$|T_{\mathrm{g}}|$为全局模型的已学习任务总数；$A_{\mathrm{Tg}}(t\to t)$为中心服务器在完成第t个训练任务后，其全局模型在第t个任务测试集上的任务准确率；$A_{\mathrm{Tg}}(|T_{\mathrm{g}}|\to t)$为在完成所有任务训练后，中心服务器的全局模型在第$t$个任务测试集上的任务准确率。

5.1.4 数据划分

为了模拟联邦学习中的客户端分布和多任务学习情况，本试验将数据集划分为4个客户端，每个客户端随机抽取数据集总数中一定比例的样本，以此模拟现实中不同客户端的数据异构性。为模拟每个客户端在联邦学习中的多任务学习情况，将每个客户端的数据进一步划分为5个子任务，每个子任务的数据集大小固定，且客户端之间的子任务具有一定的独立性。在具体训练时，客户端使用属于自己子任务的数据进行训练。训练结束后，将更新的模型权重上传至服务器进行聚合。

5.1.5 联邦环境

在本试验的联邦环境设置中，训练轮次的设定基于数据集复杂程度进行差异化安排。对于MNIST和FashionMNIST这类相对简单的数据集，因模型训练易于收敛，设置训练轮次为10轮。在这10轮中，模型有足够机会捕捉数据特征，达成较高测试准确度，稳定实现对相应数据的良好分类效果。由于SVHN和CIFAR10数据集的样本丰富、类别繁多，数据呈现较高复杂性。面对这种情况，将训练轮次提升至20轮。更多的训练轮次能够为模型提供足够时间适应不同客户端的数据特点，减少因数据差异产生的性能波动，促使模型更好地收敛，在处理复杂数据和应对异构环境时展现更强的稳定性与鲁棒性。考虑CIFAR100数据集包含更多类别和细粒度的区分要求，本研究将训练轮次进一步增加至30轮，有助于模型更充分地学习和适应这些复杂任务，提高全局模型在CIFAR-100上的表现。

5.2 对比试验

5.2.1 精准度对比

本研究在MNIST、FashionMNIST、SVHN、CIFAR10和CIFAR100数据集上，对FedAvg、FedCurv、FedHAR、FedProx、FedGReL、Powder及本研究所提MAWFCL模型进行训练。通过记录每个模型在每轮全局模型聚合后的A_{T}，分析各模型在不同任务难度下的表现。所有模型均在相同数据划分、优化器与训练轮次条件下训练，以保证对比的公平性。

MAWFCL与6种主流联邦学习方法在5个典型数据集上的A_{T}对比如表1所示。由表1可以看出，MAWFCL在多项任务中表现出良好的准确性与稳健性。在简单任务场景（MNIST和FashionMNIST）中，MAWFCL的A_{T}同其他模型基本持平。在FashionMNIST数据集上，Powder的A_{T}为90.30%，MAWFCL较Powder提升1.99百分点。这一差距可能源于Powder依赖prompt学习机制，在简单任务场景下尚未充分发挥优势。在复杂任务场景（SVHN和CIFAR10）中，MAWFCL的A_{T}分别达到84.51%和75.81%，较FedGReL分别提升

2.82百分点和 2.46 百分点,表明加权聚合策略能够有效缓解异构客户端间的知识冲突。Powder 在 SVHN 和 CIFAR10 上的表现出色,A_T 分别达到 94.02%和 78.22%,但在 CIFAR100 上的 A_T 低于 MAWFCL。在 CIFAR100 数据集上,MAWFCL 的 A_T 为 54.38%,相较 FedGReL 和 Powder 分别提升 10.93 百分点和 10.17 百分点,进一步表明在应对具有更多类别和更高复杂度的任务时,MAWFCL 凭借模块化结构和加权聚合策略展现出更强的适应性和鲁棒性。

表 1　联邦设置下全局模型的 A_T

Table 1　A_T of the global model in the federated setting

模型	A_T/%				
	MNIST	FashionMNIST	SVHN	CIFAR10	CIFAR100
FedAvg	99.30	92.25	82.51	73.83	42.95
FedCurv	99.31	92.40	—	—	—
FedHAR	99.35	92.52	—	—	—
FedProx	—	—	81.96	72.54	43.05
FedGReL	—	—	81.69	73.35	43.45
Powder	99.39	90.30	94.02	78.22	44.21
MAWFCL	99.28	92.29	84.51	75.81	54.38

注:“—”表示该方法未在对应数据集上进行试验,无法获得可用结果。

5.2.2　收敛分析

在联邦学习研究中,模型的收敛速度是评估其性能的关键指标之一。为了精准评估不同模型的收敛速度,本研究详细记录每个模型在 5 种数据集上的收敛情况,重点关注模型达到稳定准确率 A 所需的训练轮次 r。一般,收敛速度越快,意味着模型能够在较少训练轮次内实现较好的性能表现。这对于提高学习效率和资源利用效率具有重要意义。各模型在 5 种数据集上的收敛曲线如图 3 所示。

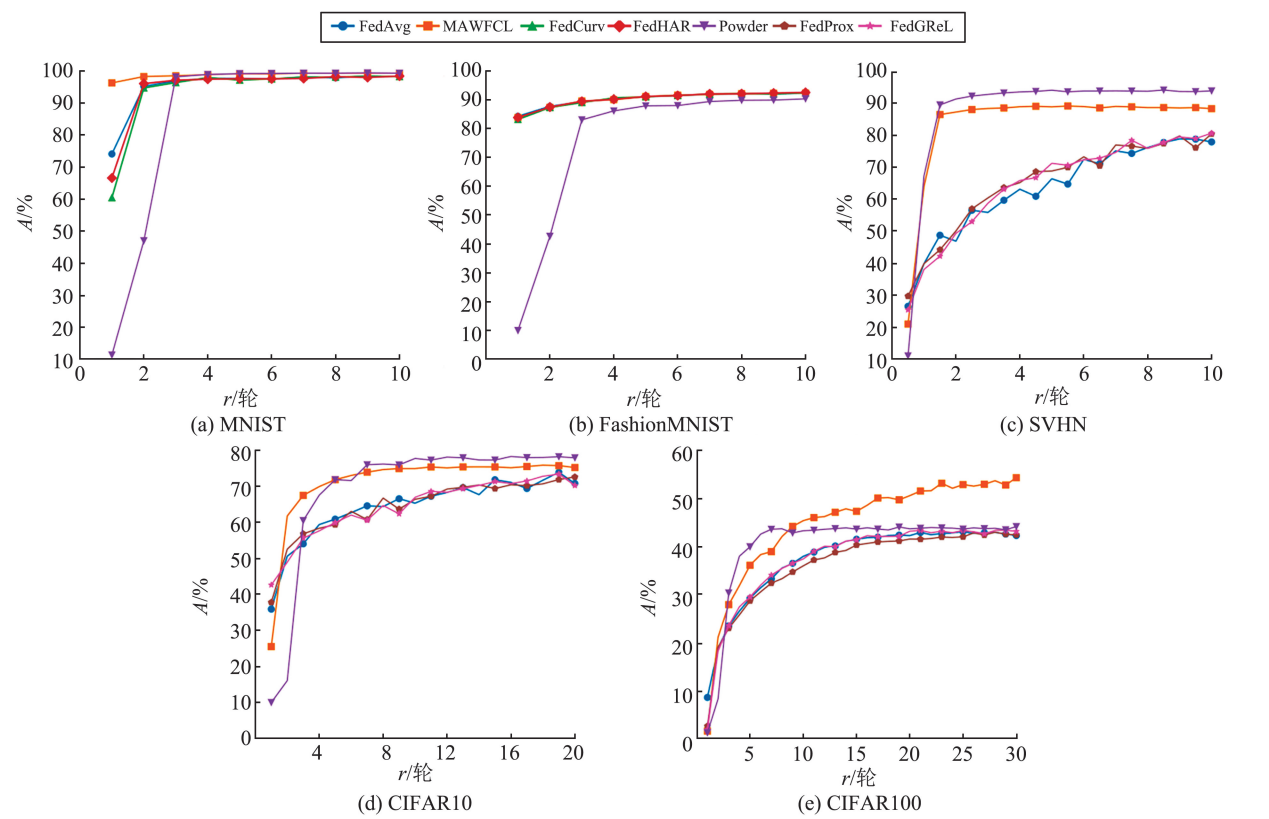

图 3　模型在 5 种数据集上的收敛曲线

Fig.3　Convergence curves of the models on five datasets

在简单任务场景(MNIST 和 FashionMNIST)中,所有模型均可较快收敛并取得较高的 A。在 MNIST 数据集中,MAWFCL 模型约在第 3 轮训练时的 A 已突破 98.0%,最终稳定在 99.2%;在 FashionMNIST 数据集中,MAWFCL 模型的 A 在第 4 轮达到 90.0%,最终稳定在 92.0%。由于数据集简

单，各模型之间的表现差距较小，但 MAWFCL 仍能在较短的训练轮次内稳定地达到较高的 A，展现出与其他模型相当的竞争力。在 FashionMNIST 数据集中，Powder 模型的收敛速度相对较慢，尽管 A 最终达到 90%左右，但仍低于 MAWFCL，差距主要在中后期逐渐显现。

在 SVHN 和 CIFAR10 中，模型间差异进一步扩大。MAWFCL 模型优势明显，在 SVHN 中，第 3 轮时 A 可达 86%左右；在 CIFAR10 中，在第 15 轮 A 达 75%左右，明显快于多数对比模型。FedProx 和 FedCurv 模型虽表现稳定，但最终 A 低于 MAWFCL 模型。综上，MAWFCL 在收敛效率和最终准确率上均表现出明显优势。Powder 在两个数据集中的 A 较高，收敛过程相对平稳，但收敛速度略慢于 MAWFCL。

在 CIFAR100 数据集上，MAWFCL 的收敛速度相对较慢，但动态聚合策略展现出强大的适应性与稳定性，最终仍取得最高 A，优于所有对比模型。

5.2.3 灾难性遗忘对比

灾难性遗忘对比试验围绕联邦学习模型在增量学习中的表现展开。试验设定 2 个顺序任务，分别基于 MNIST 和 CIFAR10 数据集，模拟先在 MNIST 上开展多轮训练及聚合，后在 CIFAR10 上重复此过程的持续学习模式。选用 4 个客户端参与联邦训练，每个客户端获取对应数据集的切分子集，于本地完成训练后，将权重上传至服务器进行全局聚合。对比模型选用 FedAvg 和 FedGReL。FedAvg 作为传统的联邦平均模型，在持续学习时容易出现遗忘问题；FedGReL 是一种带有梯度惩罚机制的联邦持续学习模型。

在训练与评估记录环节，客户端本地模型在完成每轮本地训练后，需记录在当前轮次内已学任务上的测试精度。待所有数据集任务学习完毕，基于最终的本地模型进行评估，得出对所有任务的最终精度，用于计算横向灾难性遗忘率 r_{HCF}。针对全局模型，在每完成一个任务的所有训练轮次（即 4 个客户端均上传权重并完成聚合）后，记录全局模型在所有已学任务上的测试精度。当 2 个数据集全部训练结束，统计全局模型在每个任务上的精度，计算纵向灾难性遗忘率 r_{VCF}。

灾难性遗忘对比试验结果如表 2 所示。在横向灾难性遗忘率对比中，FedAvg 模型的 r_{HCF} 为 0.46，明显高于其他模型，表明在从 MNIST 数据集任务过渡到 CIFAR10 数据集任务的学习过程中，该模型对先前在 MNIST 上所学任务的遗忘程度较为严重；FedGReL 模型的 r_{HCF} 为 0.35，略低于 FedAvg；MAWFCL 模型的 r_{HCF} 仅为 0.29，在三者中最低，充分说明 MAWFCL 模型借助自适应知识重构与模块化稀疏复用等机制，能够更好地保留先前学习任务的知识，在抑制横向灾难性遗忘方面表现卓越。在纵向灾难性遗忘率对比中，FedAvg 模型的 r_{VCF} 最高，为 0.48，意味着其全局模型在 4 个客户端本地任务上的性能差异较大，即不同客户端完成本地任务后，全局聚合模型在各客户端任务上的表现参差不齐；FedGReL 模型的 r_{VCF} 为 0.36，相对较低；MAWFCL 模型的 r_{VCF} 最低，为 0.31，体现出 MAWFCL 模型在异构环境下具备更优的泛化一致性，通过自适应平衡聚合算法等能够有效降低全局模型在不同客户端任务上的性能差异，更稳定地适应各客户端的本地任务数据分布。

表 2 联邦设置下全局模型的 r_{HCF} 和 r_{VCF}

Table 2 r_{HCF} and r_{VCF} of the global models in the federated setting

模型	r_{HCF}	r_{VCF}
FedAvg	0.46	0.48
FedGReL	0.35	0.36
MAWFCL	0.29	0.31

5.3 消融试验

5.3.1 移除模块化结构

本研究通过消融试验深入探究模块化结构对 MAWFCL 模型性能的具体影响。为确保试验结果的准确性与可靠性，除对模型结构进行特定调整外，其他试验设置均与之前的对比试验保持高度一致。移除模型结构中的本地基础参数模块和自适应参数模块，仅采用基础卷积网络开展训练。该设置聚焦于验证显式参数隔离在抑制遗忘方面发挥的关键作用，更清晰地剖析模块化结构在模型性能表现中的内在影响机制。

移除模块化结构后的模型收敛情况如图 4 所示。对比图 4 与图 3 可以发现显著差异。在 MNIST和 FashionMNIST 数据集上，移除模块化结构后，MAWFCL 模型性能明显下滑。以 FashionMNIST 为例，未移除模块化结构时，MAWFCL 模型的 A 约在第 4 轮训练时达 90.0%，最终稳定在 92.0%；移除模块化结构后，MAWFCL 模型最终的 A 约为90.1%，较未移除模块化结构时下降约 1.9 百分点，并且在第 8 轮开始出现过拟合，

曲线上升停滞并出现轻微下降趋势。面对 SVHN 和 CIFAR10 数据集,移除模块化结构对 MAWFCL 模型的影响更为显著。在 CIFAR10 数据集中,未移除模块化结构时,MAWFCL 模型在第 20 轮训练的 A 达 75.8%,移除后在第 20 轮仅为 68.8%,下降约 7.0 百分点,且在第 10 轮开始出现过拟合。在 SVHN 数据集上,相较于未移除模块化结构的模型,移除后的模型达到相近 A 的训练轮次增加,最终 A 下降约 2.5 百分点,且在第 12 轮出现明显过拟合。由此可见,模块化结构对模型应对复杂数据至关重要,可有效抑制过拟合,提升模型准确率。

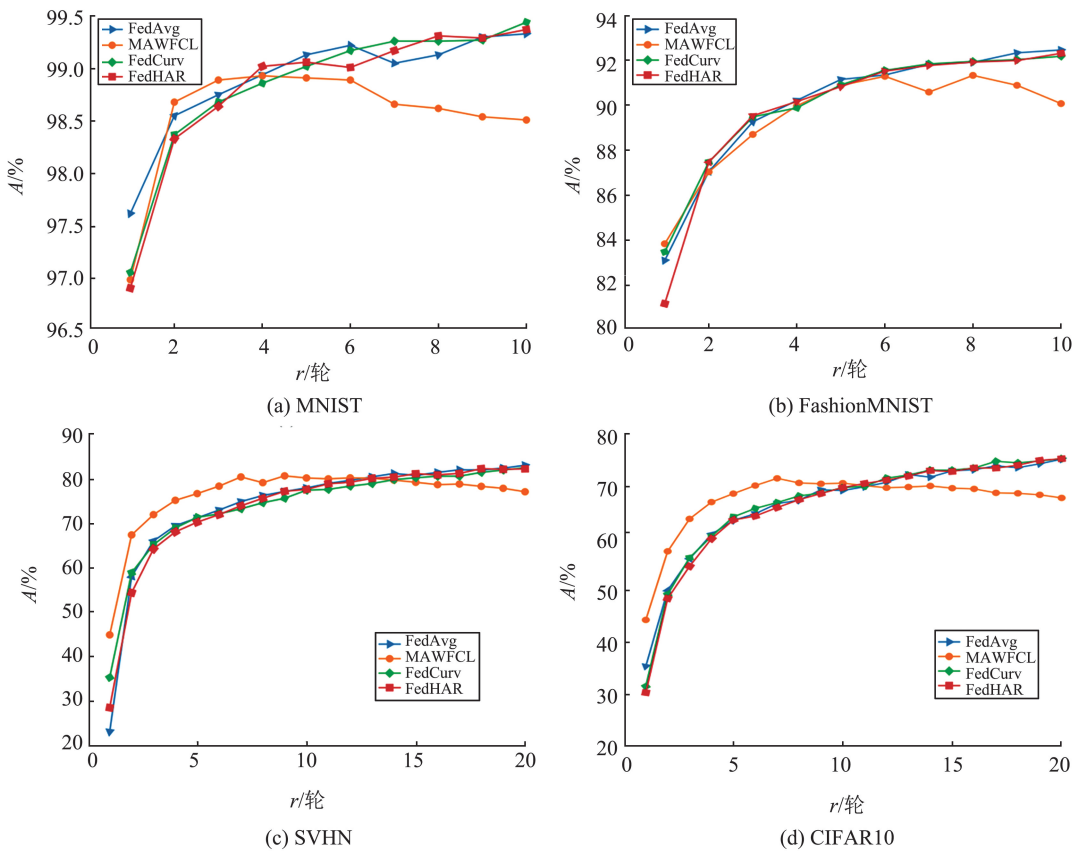

图 4　移除模块化结构后模型收敛情况

Fig.4　Model convergence status without the modular structure

5.3.2　移除精准遗忘机制

本试验通过移除模型中的精准遗忘机制对比观察模型表现,试验结果如图 5 所示。

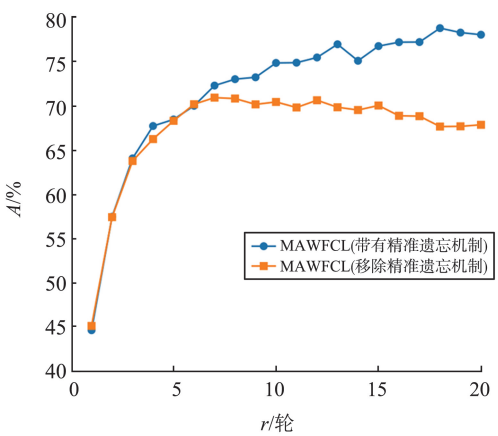

图 5　在 CIFAR10 上 MAWFCL 模型带有或移除精准遗忘机制的性能对比

Fig.5　Performance comparison of the MAWFCL model with or without the precise forgetting mechanism on CIFAR10

由图 5 可以看出:去除该机制后,MAWFCL 在多任务学习场景中的灾难性遗忘问题加剧,特别是在引入新任务时,MAWFCL 在之前已学习任务上的 A 出现显著下滑;在训练过程中,随着训练轮次增加,带有精准遗忘机制模型的 A 逐步提升并维持在相对较高水平,移除精准遗忘机制模型的 A 提升相对缓慢且后期有下降趋势。

带有精准遗忘机制模型和移除精准遗忘机制模型的 r_{HCF} 和 r_{VCF} 对比结果如表 3 所示。由表 3 可以看出,带有精准遗忘机制模型的 r_{HCF} 和 r_{VCF} 均小于移除该机制的模型,表明精准遗忘机制能有效降低 MAWFCL 在多任务学习中的灾难性遗忘程度。无论是在客户端任务层面还是全局模型层面,精准遗忘机制都有助于提升模型对先前任务知识的保留能力和在异构环境下的泛化一致性。

表 3 带有或移除精准遗忘机制的 r_{HCF} 和 r_{VCF}
Table 3 Comparison of r_{HCF} and r_{VCF} with or without the precise forgetting mechanism

MAWFCL 设置	r_{HCF}	r_{VCF}
带有精准遗忘机制	0.29	0.31
移除精准遗忘机制	0.46	0.48

5.4 通信效率评估

本研究在 CIFAR10 数据集中,对 FedAvg、FedGReL 和 MAWFCL 模型的通信效率展开评估,评估指标包含收敛轮数、单轮通信量及总通信成本。收敛轮数指模型达到目标准确率(CIFAR10 为 80%)时所需的全局聚合轮数,轮数越少,表明模型的通信效率越高;单轮通信量是客户端上传的模型参数总量,用于反映带宽压力;总通信成本通过收敛轮数与单轮通信量相乘得出,综合评估模型的整体通信效率。

模型通信效率评估结果如表 4 所示。由表 4 可以看出:MAWFCL 模型的收敛速度更快,仅需 32 轮即可达到目标准确率;相比 FedGReL,MAWFCL 的总通信成本降低 5.7%;尽管 MAWFCL 的单轮通信量高于 FedAvg,但凭借动态聚合策略减少了冗余传输,有效将总通信成本控制在合理范围内。上述结果表明,MAWFCL 在通信效率方面具有较好的综合表现。

表 4 模型通信效率评估
Table 4 Evaluation of model communication efficiency

模型	收敛轮数/轮	单轮通信量/kB	总通信成本/kB
FedAvg	45	12.4	558.0
FedGReL	38	18.3	695.4
MAWFCL	32	20.5	656.0

6 结论

本研究针对联邦持续学习中因任务异构性和数据非独立同分布引发的横向与纵向灾难性遗忘问题,提出一种基于模块化网络的自适应加权联邦持续学习方法。通过构建共享基础知识参数与客户端自适应参数的模块化结构,结合自适应知识重构过程、模块化稀疏复用策略、自适应平衡聚合算法及精准遗忘机制,MAWFCL 实现新任务学习与历史知识保持之间的有效平衡。理论分析表明,MAWFCL 在参数隔离与动态聚合方面具有良好的数学表征和优化意义,能够在资源受限的联邦学习场景下有效缓解灾难性遗忘现象。在 MNIST、FashionMNIST、SVHN、CIFAR10 和 CIFAR100 等数据集上的试验结果显示,MAWFCL 在测试准确率、收敛速度和灾难性遗忘控制方面优于现有主流方法。在简单任务场景下,MAWFCL 表现稳健;在 CIFAR100 等复杂任务上,MAWFCL 展现出显著优势。这充分证明 MAWFCL 在异构任务和多数据源场景下具有更强的适应性和鲁棒性。

尽管 MAWFCL 取得了较好的试验结果,但现有研究仍存在一定局限。未来工作将围绕以下几个方向展开:当前随机划分策略难以完全模拟真实环境中的分布差异,未来将引入基于 Dirichlet 分布的 Non-IID 数据划分机制,进一步评估 MAWFCL 方法在不同异构性条件下的泛化性和稳定性;扩展 MAWFCL 方法至多任务学习、序列建模等异构任务场景,评估其在不同任务结构和输入空间下的适应性和泛化能力;现阶段的模块选择策略主要依赖相似度阈值和容量约束等启发式规则,未来将探索任务感知的模块调度机制,引入任务特定的识别与评分机制,提升模块复用的精细度与系统整体效率。综上,MAWFCL 为应对联邦持续学习中的任务异构与知识冲突问题提供一种可扩展、可解释且训练友好的解决方案。通过进一步增强对数据分布、任务结构和模块调度的自适应性,期望能够提升 MAWFCL 方法在大规模联邦持续学习系统中的应用效果,拓展其在复杂场景中的适用性。

参考文献:

[1] WANG L Y, ZHANG X X, SU H, et al. A comprehensive survey of continual learning: theory, method and application[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(8): 5362-5383.

[2] 王文晟, 谭宁, 黄凯, 等. 基于大模型的具身智能系统综述[J]. 自动化学报, 2025, 51(1): 1-19.
WANG Wensheng, TAN Ning, HUANG Kai, et al. Embodied intelligence systems based on large models: a survey[J]. Acta Automatica Sinica, 2025, 51(1): 1-19.

[3] MCLEAN S, READ G J M, THOMPSON J, et al. The risks associated with artificial general intelligence: a systematic review[J]. Journal of Experimental & Theoretical Artificial Intelligence, 2023, 35(5): 649-663.

[4] SILVER D, HUBERT T, SCHRITTWIESER J, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play[J]. Science, 2018, 362(6419): 1140-1144.

[5] 肖雄, 唐卓, 肖斌, 等. 联邦学习的隐私保护与安全防御研究综述[J]. 计算机学报, 2023, 46(5):

1019-1044.
XIAO Xiong, TANG Zhuo, XIAO Bin, et al. A survey on privacy protection and security defense in federated learning[J]. Chinese Journal of Computers, 2023, 46(5): 1019-1044.

[6] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]// Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS). Fort Lauderdale, USA: JMLR, 2017: 1273-1282.

[7] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks[EB/OL]. (2020-04-21)[2025-05-22]. https://arxiv.org/abs/1812.06127

[8] YU H, YANG X, GAO X, et al. Personalized federated continual learning via multi-granularity prompt[C]// Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona, Spain: ACM, 2024: 4023-4034.

[9] 杜甜, 陈星延, 寇纲, 等. 面向云边个性化模型解耦的聚类联邦学习方法[J]. 计算机学报, 2025, 48(2): 407-432.
DU Tian, CHEN Xingyan, KOU Gang, et al. Clustered federated learning with cloud-edge personalized model decoupling[J]. Chinese Journal of Computers, 2025, 48(2): 407-432.

[10] DE LANGE M, ALJUNDI R, MASANA M, et al. A continual learning survey: defying forgetting in classification tasks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(7): 3366-3385.

[11] YOON J, JEONG W, LEE G, et al. Federated continual learning with weighted inter-client transfer[C]// Proceedings of the 38th International Conference on Machine Learning (ICML). [S. l.]: PMLR, 2021: 12073-12086.

[12] YU H Z, CHEN Z K, ZHANG X, et al. FedHAR: semi-supervised online learning for personalized federated human activity recognition[J]. IEEE Transactions on Mobile Computing, 2023, 22(6): 3318-3332.

[13] 王鹏飞, 魏宗正, 周东生, 等. 联邦忘却学习研究综述[J]. 计算机学报, 2024, 47(2): 396-422.
WANG Pengfei, WEI Zongzheng, ZHOU Dongsheng, et al. A survey on federated unlearning[J]. Chinese Journal of Computers, 2024, 47(2): 396-422.

[14] WUERKAIXI A, CUI S, ZHANG J F, et al. Accurate forgetting for heterogeneous federated continual learning[EB/OL]. (2025-02-20)[2025-05-22]. https://arxiv.org/abs/2502.14205

[15] YOON J, KIM S, YANG E, et al. Scalable and order-robust continual learning with additive parameter decomposition[EB/OL]. (2020-02-15)[2025-05-22]. https://arxiv.org/abs/1902.09432

[16] MALLYA A, LAZEBNIK S. PackNet: adding multiple tasks to a single network by iterative pruning[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 7765-7773.

[17] KIRKPATRICK J, PASCANU R, RABINOWITZ N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(13): 3521-3526.

[18] LI Z Z, HOIEM D. Learning without forgetting[C]// European Conference on Computer Vision (ECCV). Amsterdam, Netherlands: Springer, 2016: 614-629.

[19] CHAUDHRY A, RANZATO M, ROHRBACH M, et al. Efficient lifelong learning with A-GEM[EB/OL]. (2019-01-09)[2025-05-22]. https://arxiv.org/abs/1812.00420

[20] YOON J, YANG E, LEE J, et al. Lifelong learning with dynamically expandable networks[EB/OL]. (2018-06-11)[2025-05-22]. https://arxiv.org/abs/1708.01547

[21] ALJUNDI R, CHAKRAVARTY P, TUYTELAARS T. Expert gate: lifelong learning with a network of experts[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 7120-7129.

[22] ROLNICK D, AHUJA A, SCHWARZ J, et al. Experience replay for continual learning[EB/OL]. (2019-11-26)[2025-05-22]. https://arxiv.org/abs/1811.11682

[23] BUZZEGA P, BOSCHINI M, PORRELLO A, et al. Dark experience for general continual learning[C]// Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: ACM, 2020: 15920-15930.

[24] REBUFFI S A, KOLESNIKOV A, SPERL G, et al. iCaRL: Incremental classifier and representation learning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017: 5533-5542.

[25] ZHU Z P, ZHAO S J, CHU C C, et al. FedPMR: federated personalized mixture representation for driver intention prediction[J]. IEEE Transactions on Intelligent Vehicles, 2025, 10(1): 627-640.

[26] HE Y T, CHEN Y Q, YANG X D, et al. Class-wise

adaptive self-distillation for federated learning on Non-IID data[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(11): 12967-12968.

[27] 姜慧, 何天流, 刘敏, 等. 面向异构流式数据的高性能联邦持续学习算法[J]. 通信学报, 2023, 44(5): 123-136.
JIANG Hui, HE Tianliu, LIU Min, et al. High-performance federated continual learning algorithm for heterogeneous streaming data[J]. Journal on Communications, 2023, 44(5): 123-136.

[28] DONG J H, WANG L X, FANG Z, et al. Federated class-incremental learning[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE, 2022: 10154-10163.

[29] SAHA G, GARG I, ROY K. Gradient projection memory for continual learning[EB/OL]. (2021-03-17)[2025-05-22]. https://arxiv.org/abs/2103.09762

[30] LUO P Y X, HAN R, ZHANG Q L, et al. FedKNOW: federated continual learning with signature task knowledge integration at edge[C]//2023 IEEE 39th International Conference on Data Engineering (ICDE). Anaheim, USA: IEEE, 2023: 341-354.

[31] SHOHAM N, AVIDOR T, KEREN A, et al. Overcoming forgetting in federated learning on Non-IID data[EB/OL]. (2019-10-17)[2025-05-22]. https://arxiv.org/abs/1910.07796

[32] WU C H, WU F Z, QI T, et al. FedCL: federated contrastive learning for privacy-preserving recommendation[EB/OL]. (2022-04-21)[2025-05-22]. https://arxiv.org/abs/2204.09850

[33] PIAO H M, WU Y C, WU D P, et al. Federated continual learning via prompt-based dual knowledge transfer[C]// Proceedings of the 41st International Conference on Machine Learning. Vienna, Austria: JMLR, 2024: 40725-40739.

[34] WANG Q, LIU B Y, LI Y W. Traceable federated continual learning[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2024: 12872-12881.

[35] HE Y C, SHEN C Y, WANG X F, et al. FPPL: an efficient and Non-IID robust federated continual learning framework[C]//2024 IEEE International Conference on Big Data (BigData). Washington, DC, USA: IEEE, 2024: 3692-3701.

[36] YU H, YANG X, GAO X, et al. Personalized federated continual learning via multi granularity prompt[C]// Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona, Spain: ACM, 2024: 4023-4034.

[37] WANG Z R, DAI Z H, PÓCZOS B, et al. Characterizing and avoiding negative transfer[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2020: 11285-11294.

(编辑:孙亚彤)

(上接第 18 页)

[15] CHEN T Q, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA: ACM, 2016: 785-794.

[16] ZHUO H, LI T R, LU W, et al. Prediction model for spontaneous combustion temperature of coal based on PSO-XGBoost algorithm[J]. Scientific Reports, 2025, 15(1): 2752.

[17] GUPTA A, GOWDA S, TIWARI A, et al. XGBoost-SHAP framework for asphalt pavement condition evaluation[J]. Construction and Building Materials, 2024, 426: 136182.

(编辑:孙亚彤)